

# Classificação das Configurações de Mão da Língua Brasileira de Sinais Mediante Rede Neural Artificial Kohonen

## Classification of Hand Configurations of the Brazilian Sign Language Using Neural Network Kohonen

Kelly Lais Wiggers

Universidade Estadual do Centro-Oeste - UNICENTRO, Guarapuava, PR

*kellylaiswiggers@hotmail.com*

Angelita Maria de Ré

Departamento de Ciência da Computação

Universidade Estadual do Centro-Oeste - UNICENTRO, Guarapuava, PR

*amdre65@yahoo.com.br*

Andres Jessé Porfírio

BlackFace - Núcleo de Pesquisas no Ensino da Computação

Universidade Tecnológica Federal do Paraná - UTFPR, Guarapuava, PR

*andresporfrio@utfpr.edu.br*

**Resumo:** A falta de assistência tecnológica para a inclusão do surdo na sociedade contemporânea é uma realidade existente. Este trabalho consiste em realizar a classificação das configurações de mão da Língua Brasileira de Sinais (LIBRAS), mediante Rede Neural Artificial Kohonen, que se baseia em aprendizado competitivo, simulando processos específicos do cérebro humano. O sistema é composto pelas seguintes etapas: extração de características das configurações de mão (imagens), treinamento da Rede Neural Artificial, testes e análise dos resultados.

**Palavras-chave:** configurações de mão de LIBRAS; extração de características de imagens; mapas auto-organizáveis de Kohonen.

Recebido em 09/10/2013 - Aceito em 11/08/2014.

---

RECEN	16(2)	p. 175-197	jul/dez	2014	DOI: 10.5935/RECEN.2014.02.02
-------	-------	------------	---------	------	-------------------------------

---

**Abstract:** The lack of technological support for the inclusion of the deaf in contemporary society is an existing reality. This work consists of performing the classification of hand configurations of the Brazilian Sign Language (LIBRAS) using Kohonen Artificial Neural Network, which is based on competitive learning, simulating specific processes in the human brain. The system consists of the following steps: feature extraction of hand configurations (images), Artificial Neural Network training, testing and computational results.

**Key words:** extracting features of images; hand configurations of the Brazilian sign language; Kohonen artificial neural network.

## 1 Introdução

Os sistemas computacionais estão presentes em diversas áreas de atuação profissional, servindo como uma ferramenta de grande importância na coleta, armazenamento e análise de informações. No reconhecimento dos sinais na Língua Brasileira de Sinais isso também é uma realidade.

A Língua Brasileira de Sinais é a principal forma de comunicação utilizada pela comunidade surda brasileira. Para suprir as necessidades de tecnologia desta comunidade, estão sendo desenvolvidos sistemas como o SIGUS [1] e o F-LIBRAS [2]. O principal objetivo desses sistemas é oferecer subsídios tecnológicos a fim de favorecer os deficientes auditivos na comunicação.

Para reconhecer o sinal da LIBRAS, são necessários diversos parâmetros tais como: configuração de mão, movimento e expressão facial, dentre outros. Essa grande quantidade de informações e a maneira em que se relacionam, dificulta o desenvolvimento de um sistema computacional capaz de efetuar o reconhecimento eficiente da língua de sinais [3].

Dentre as diversas técnicas possíveis para classificação das configurações de mão em LIBRAS estão as Redes Neurais Artificiais, que são sistemas computacionais com o objetivo de solucionar problemas através da simulação do modo de funcionamento dos neurônios, o que pode ser utilizado para a percepção das configurações de mão da língua de sinais.

A Rede Neural Artificial Kohonen tem sido frequentemente utilizada em inúmeras aplicações, incluindo áreas como reconhecimento de padrões, processamento de sinais, compressão de dados, modelagem biológica e mineração de dados. A Kohonen é um tipo de Rede Neural Artificial baseada em aprendizado competitivo e não supervisionado. Propriedades importantes como a capacidade de aproximar o espaço de entrada, ordenação topológica e casamento de densidade, aliadas, a simplicidade do modelo e a facilidade de implementação do seu algoritmo de aprendizagem justificam o sucesso da Kohonen [4].

A necessidade de recursos tecnológicos da comunidade surda brasileira associada à inexistência de sistemas funcionais e acessíveis para atender suas necessidades específicas foram os propulsores do desenvolvimento deste trabalho. Portanto, o objetivo deste trabalho é realizar a classificação das configurações de mão de LIBRAS por Rede Neural Artificial Kohonen, a partir de uma Base de Dados criada por um grupo de pesquisa da Universidade Federal do Paraná.

## **2 Redes neurais artificiais**

As Redes Neurais Artificiais (RNA) são um instrumento desenvolvido na década de 40, por Pitts e McCulloch. A ideia principal fazia uma analogia entre os neurônios biológicos e circuitos eletrônicos, capazes de simular conexões sinápticas pelo uso de resistores variáveis e amplificadores [5].

Os trabalhos em RNAs têm sido motivados desde o começo pelo reconhecimento de que o cérebro humano processa informações de uma forma inteiramente diferente do computador digital convencional. O cérebro é um computador altamente complexo, não linear e paralelo. Ele tem a capacidade de organizar seus constituintes estruturais, conhecidos por neurônios, de forma a realizar certos processamentos muito mais rapidamente de que o mais rápido computador digital existente [6]. Sendo assim, a RNA é “uma rede massivamente paralela interconectada de elementos e suas organizações hierárquicas que estão intencionadas para interagir com objetos do mundo real do mesmo modo que um sistema nervoso biológico faz” [7].

As RNAs são capazes de resolver, basicamente, problemas de aproximação, predição, classificação, categorização e otimização. Aplicações conhecidas de RNAs como

reconhecimento de caracteres, reconhecimento de voz, predição de séries temporais, são alguns exemplos. Em processamento de sinais e imagens, as RNAs ganharam grande popularidade nos últimos anos como ferramenta alternativa e às vezes complementar as técnicas clássicas. Nestas áreas, as aplicações das RNAs envolvem reconhecimento de caracteres escritos a mão, compactação de dados, codificação, reconhecimento da fala, análise de componentes independentes, entre outras [8].

A capacidade de uma Rede Neural Artificial realizar o reconhecimento de padrões provém de boa organização da rede, bem como sua estrutura, resultando na etapa de treinamento bem sucedida. Diversas aplicações de Redes Neurais Artificiais dizem respeito à classificação de padrões. Nestas aplicações, a RNA aprende a classificar um padrão desconhecido entre várias classes ou categorias. Assim, ao deparar com um padrão desconhecido, deve classificá-lo em uma das classes existentes [9]. Como exemplos, têm-se os artigos de Carneiro [10] e Lamar [11], que trabalham com classificações de padrões relacionados às posturas manuais, demonstrando a viabilidade das RNAs para este tipo de problema. Estas posturas se referem ao gestos formadores do alfabeto das línguas de sinais, como por exemplo o alfabeto da Libras, que utiliza 20 configurações de mão estáticas, ou seja, que não exigem movimentos durante a realização do sinal. Já o presente trabalho utilizará todas as configurações de mão estáticas, visto que muitos sinais de Libras podem ser obtidos por mais de uma configuração de mão.

### 3 Rede Kohonen

A rede Kohonen é baseada em mapas que preservam a topologia da auto-organização. Dessa maneira, as entradas “se tornam especificamente ajustadas a vários padrões ou classes de padrões através de um processo de aprendizagem não supervisionado. Os resultados de aprendizagem encontrados parecem muito naturais, remontando um processo adaptativo similar ao encontrado no cérebro humano, o que constitui justificativa suficiente para chamar estes mapas de redes neurais” [12].

Em um mapa auto-organizável Kohonen, os neurônios estão colocados em nós de uma grade que normalmente são uni ou bidimensionais. Os neurônios se tornam seletivamente sintonizados a vários padrões de entrada ou classe de padrões de entrada

no decorrer de um processo de aprendizagem [6].

Uma das aplicações mais conhecidas da rede Kohonen é o trabalho para o reconhecimento de fonemas em finlandês utilizando mapas auto organizáveis com treinamento competitivo [9]. Este sistema classifica com sucesso 21 fonemas. O processo de reconhecimento é dividido em duas etapas: a primeira envolve a categorização dos fonemas, e a segunda, a classificação propriamente dita. A primeira etapa organiza os dados (fonemas) em categorias com o propósito de serem classificados. Na etapa de classificação, cada fonema ativa uma região do mapa, e uma sequência de fonemas correspondente a um sinal de voz ativa regiões distintas em sequência, descrevendo então uma trajetória no mapa. Esta trajetória é responsável pelo processo de reconhecimento.

Com base neste exemplo, de classificação de fonemas somado à capacidade da rede Kohonen de realizar um aprendizado não supervisionado de forma competitiva —simulando processos específicos de aprendizagem do cérebro humano—, foi escolhida esta rede para o desenvolvimento do sistema computacional para classificação das configurações de mão.

### 3.1 O algoritmo da rede Kohonen

O algoritmo responsável pela organização da rede inicia o processo arbitrando pequenos valores aleatórios aos pesos sinápticos, para que nenhuma organização prévia seja imposta ao mapa. Em seguida, dar-se-ão os processos de competição, cooperação e adaptação sináptica [5].

No Processo Competitivo, considera-se um vetor de entrada, representado por  $X_i = [X_1, X_2, \dots, X_n]^T$ , sendo um dos exemplos de treinamento, é apresentado à rede sem que se especifique a saída desejada. Um neurônio  $y$  deverá responder melhor a esta entrada, ou seja, será o neurônio vencedor. O critério para a escolha do neurônio vencedor pode se basear em qualquer métrica que informe a distância entre o vetor de entrada e o vetor de pesos dos neurônios. É frequente o uso da distância euclidiana. Assim, o neurônio que apresenta a menor distância euclidiana entre o vetor de entrada e o seu vetor de pesos é o vencedor.

Considera-se que o vetor  $W_{ij} = [W_{i1}, W_{i2}, \dots, W_{in}]^T$  represente os valores dos

pesos sinápticos de um neurônio  $i$  no instante  $t$  e o vetor  $X_i = [X_1, X_2, \dots, X_n]^T$ , represente um exemplo de treinamento apresentado à rede no mesmo instante  $t$ , a distância euclidiana entre  $W_i$  e  $X_i$ , será dada pela expressão (1):

$$d_i(t) = a_0 + \sum_{j=1}^N x_j(t) - w_{ij}(t)^2 \quad (1)$$

em que:

- $d_i(t)$  = distância euclidiana entre o vetor de pesos do neurônio  $i$  e o vetor de entrada, na iteração  $t$ ;
- $i$  = índice do neurônio;
- $j$  = índice do nó de entrada;
- $N$  = número de sinais de entrada (números de dimensões do vetor  $x$ );
- $X_j(t)$  = sinal de entrada do nó  $j$  na iteração  $t$ ;
- $W_{ij}(t)$  = valor do peso sináptico entre o nó de entrada  $j$  e o neurônio  $i$  na iteração  $t$ .

No Processo Cooperativo, o neurônio vencedor determina o centro de uma vizinhança topológica de neurônios cooperativos. A definição desta vizinhança topológica é baseada na evidência neurobiológica de que há interação lateral entre um conjunto de neurônios biológicos excitados. Em particular, um neurônio classificado como vencedor tende a excitar mais os neurônios em sua vizinhança imediata, do que aqueles neurônios que estão mais distantes. Assim, pode-se afirmar que a vizinhança topológica ao redor do neurônio vencedor  $i$  decresce suavemente com a distância lateral [13].

Seja  $h_{ij}$  a vizinhança topológica centrada no neurônio vencedor  $i$  e circundada por um conjunto de neurônios excitados cooperativos, dos quais um neurônio típico é denotado por  $j$ . Seja, ainda,  $d_{ij}$  a distância lateral entre o neurônio vencedor  $i$  e o neurônio excitado  $j$ . Pode-se então assumir que a vizinhança topológica  $h_{ij}$  é uma função unimodal (que apresenta um único mínimo ou máximo) da distância lateral

$d_{ij}$  (relativa a distância topológica dos neurônios), tal que satisfaça duas exigências distintas [14]:

- (i) A vizinhança topológica  $h_{ij}$  é simétrica ao redor do ponto máximo definido por  $d_{ij} = 0$ , ou seja, atinge seu valor máximo no neurônio vencedor  $i$ , para o qual a distância  $d_{ij}$  é zero;
- (ii) A amplitude da vizinhança topológica  $h_{ij}$  decresce monotonamente com o aumento da distância lateral  $d_{ij}$ , decaindo a zero para  $d_{ij} \rightarrow \infty$ ; condição necessária para convergência.

Várias funções satisfazem as exigências para a vizinhança topológica  $h_{ij}$  e são apresentadas em (2) [13]:

$$\begin{aligned} h_{ji}(t) &= (\alpha_t - d_{ij}) \\ h_{ji}(t) &= e^{\frac{-d_{ji}^2}{2\alpha_t^2}} \\ h_{ji}(t) &= e^{\frac{-d_{ji}^2}{2\alpha_t^2}} (\alpha_t - d_{ji}) \end{aligned} \tag{2}$$

$$h_{ji}(t) = \text{Max} \left\{ 0, 1 - (\alpha_t - d_{ji})^2 \right\}$$

em que  $\alpha_t$  é o raio da vizinhança topológica no tempo  $t$  ( $t$  representa o número de iterações), mede o grau com que os neurônios excitados, vizinhos do neurônio vencedor, participam do processo de aprendizagem.

O parâmetro  $\alpha_t$  deve diminuir com o passar do tempo, o que implica em valores de  $h_{ij}$  menores ao longo do tempo, caracterizando uma vizinhança mais restrita, e, portanto, mais especializada [5]. A mais utilizada é a função de vizinhança gaussiana das equações de (2), a qual é invariante à translação, isto é, independente da localização do neurônio vencedor [14].

No Processo Adaptativo, considerando-se um peso sináptico  $W_{ij}$  entre o nó de entrada  $j$  e o neurônio  $i$ , o ajuste  $\Delta W_{ij}$  que deve sofrer, será regido pela Equação (3)

[5]:

$$\Delta W_{ij} = (t) \cdot h_{ij}(t) \cdot (X_j - W_{ij}) \quad (3)$$

em que o termo  $h_{ij}(t)$  é o parâmetro vizinhança topológica na iteração  $t$ , no qual o índice  $k$  se refere ao neurônio melhor classificado  $k$ .

O parâmetro taxa de aprendizagem,  $\eta(t)$ , geralmente é definido pela expressão em (4):

$$\eta(t) = \eta_0 \cdot e^{-\frac{t}{t1}} \quad (4)$$

em que  $t1$  é uma constante de tempo e  $\eta_0$  é o valor inicial adotado, o qual deve ser maior que zero e menor que um.

Portanto, dada uma rede de Kohonen, o algoritmo para treiná-la é descrito pelas etapas de definição aleatória dos pesos iniciais dos arcos, definição de  $\alpha_0$ ,  $\eta_0$ ,  $t_1$  e o valor do erro, inserção de exemplo de treinamento e cálculo da distância euclidiana para ser encontrado o neurônio vencedor, respectivamente. Após isso, para cada neurônio é necessário calcular a largura efetiva  $\alpha_t$ , a distância lateral  $d_{ij}$ , a taxa de aprendizagem  $\eta_t$  e ajustar os pesos dos arcos [5].

## 4 Processamento de imagens

Os problemas de visão computacional e reconhecimento de imagens possuem características importantes, e a principal delas é certamente a quantidade de informação manipulada. Em reconhecimento de imagens, os problemas requerem o tratamento de imagens de alta definição com uma grande quantidade de informação. O tratamento destes problemas com RNAs frequentemente requer a utilização de técnicas convencionais de processamento de imagens para pré-processar os dados e, então, apresentá-los de uma maneira tratável para a rede [9].

Processamento de imagens é um processo onde a entrada do sistema é uma imagem e a saída é um conjunto de valores numéricos, que podem ou não compor outra imagem. A visão computacional procura emular a visão humana, portanto também possui como entrada uma imagem, porém, a saída é uma interpretação da imagem



como um todo, ou parcialmente [15].

O desenvolvimento de sistemas de visão computacional requer uma entrada de dados (imagem) geralmente obtida de sensores, câmeras ou vídeos. Essas imagens são processadas e transformadas em uma informação esperada. O processo de transformação da imagem pode ser realizado por métodos contidos em bibliotecas de processamento gráfico. Uma das bibliotecas mais famosas de processamento de imagens é o *OpenCV* [16].

A segmentação de imagens consiste em separar o conteúdo de interesse na imagem. Algoritmos de segmentação permitem achar diferenças entre dois ou mais objetos, e distinguir as partículas umas das outras e do fundo. Esta distinção permitirá ao programa interpretar pixels contíguos e agrupá-los em regiões. Os algoritmos de segmentação para imagens monocromáticas são geralmente baseados em uma das seguintes propriedades básicas de valores de níveis de cinza: descontinuidade e similaridade. Na descontinuidade a abordagem é particionar a imagem baseado em mudanças bruscas nos níveis de cinza. As principais áreas de interesse são a detecção de pontos isolados, detecção de linhas e bordas na imagem. Na similaridade, as principais abordagens baseiam-se em limiarização e crescimento de regiões [17].

O método de limiarização (*thresholding*), “consiste basicamente em separar os grupos de cinza de uma imagem. Em outras palavras, determina-se uma intensidade de cinza, chamada de limiar, que separa as partes da imagem, de forma que grupos de pixels com intensidades parecidas sejam separados de outros”. O principal complicador deste processo é determinar qual é o valor do limiar. De forma geral, o que pode ser feito é um histograma com as intensidades dos pixels da imagem e a quantidade de pixels com cada intensidade. Com isto, pode-se observar a formação de vales na intensidade. O valor de limiar é escolhido neste vale, como sendo o mais próximo ao meio da escala dos níveis de cinza [18].

Os histogramas são ferramentas de processamento de imagens que possuem grande aplicação prática. São determinados a partir de valores de intensidade dos pixels. Entre as principais aplicações dos histogramas estão a melhora da definição, a compressão de imagens, a segmentação de imagens ou ainda a descrição de uma imagem [15].

Estes métodos de processamento de imagens explicados serão utilizados em ima-

gens das configurações de mão da Língua Brasileira de Sinais, melhor detalhado nas Seções 5 e 6.

## 5 Língua brasileira de sinais (Libras)

A Língua Brasileira de Sinais - LIBRAS - é a forma de comunicação oficial utilizada pela comunidade surda brasileira. Trata-se de uma língua natural de sinais, completamente estruturada, com próprias e bem definidas regras morfológicas, sintáticas e semânticas, com origem na língua de sinais francesa [19].

A LIBRAS não pode ser estudada tendo como base a Língua Portuguesa, porque ela tem gramática diferenciada, independente da língua oral. A ordem dos sinais na construção de um enunciado obedece a regras próprias que refletem a forma de o surdo processar suas ideias, com base em sua percepção visual-espacial da realidade [20].

Pelo fato das línguas de sinais não serem meras transcrições de línguas faladas, é compreensível a dificuldade de comunicação existente entre pessoas não surdas e surdas, mesmo por meio da escrita, uma vez que pessoas surdas podem não ser alfabetizadas em Português [19].

Para efetuar um reconhecimento de qualquer sinal em LIBRAS, é necessário observar os seguintes parâmetros: Configuração de mão, Movimento, Expressão não manual, Ponto de articulação, Orientação da palma da mão.

As Configurações de Mão são as diversas formas que uma ou as duas mãos tomam na realização do sinal. Movimento é um parâmetro tão complexo que pode envolver uma grande quantidade de formas e direções, desde os movimentos internos da mão, os movimentos do pulso, movimentos direcionais no espaço e até conjuntos de movimentos no mesmo sinal. Ponto de Articulação é o espaço em frente ao corpo ou uma região do próprio corpo, onde os sinais são articulados. Orientação da mão é a direção para a qual a palma da mão aponta na produção do sinal. Expressões não manuais (expressão facial, dos olhos, da cabeça ou do tronco) prestam-se a dois papéis nas línguas de sinais, sendo marcação de construções sintáticas e diferenciação de itens lexicais [21].

O foco do trabalho são as configurações de mão. Estas configurações são o pa-

râmetro primitivo para realização de qualquer sinal da LIBRAS. Pode ser a mesma ou variar para outra configuração de mão durante a execução de um sinal, ou seja, existem sinais que são constituídos por mais de uma configuração de mão [22].



Figura 1. Imagens das 61 configurações de mão da LIBRAS <sup>1</sup>.

Pode haver variações no quadro de configurações, dependendo do autor. Portanto, é sempre interessante analisar qual destes é relevante para ser utilizado. A figura 1 representa um quadro composto por 61 configurações de mão da LIBRAS utilizadas para o desenvolvimento do sistema computacional proposto.

## 6 Material e métodos

As imagens das configurações de mão foram armazenadas em uma Base de Dados. Esta Base foi criada por um grupo de pesquisas da Universidade Federal do Paraná (UFPR), e compõe-se de 9 pastas. Cada pasta contém 61 imagens, sendo uma de cada configuração de mão. Destaca-se que na base de imagens utilizada neste trabalho não constam informações sobre os métodos de aquisição de imagens, como por exemplo,

<sup>1</sup><http://www.pead.faced.ufrgs.br/sites/publico/eixo7/libras/unidade4/unidade4.htm>

cor de fundo, iluminação, distância da câmera ou posicionamento de mãos. Dessa forma, dificultando a interferência no tratamento das imagens.

A figura 2 mostra uma visão geral do desenvolvimento do trabalho, com as respectivas etapas realizadas desde a obtenção da imagem até a saída classificada.

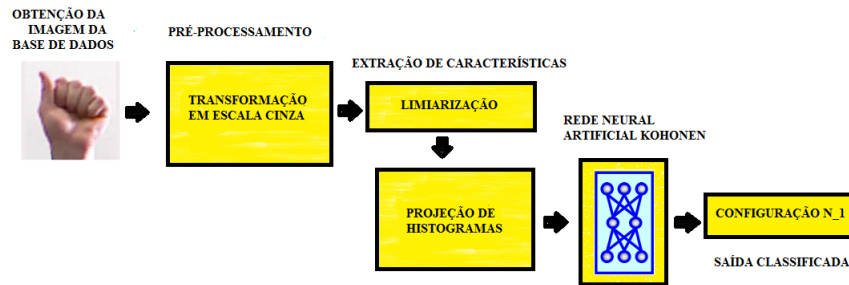


Figura 2. Visão geral do desenvolvimento do trabalho.

## 6.1 Processamento das imagens da Base de Dados

A imagem é obtida a partir da Base de Dados. A etapa de pré-processamento prepara a imagem para ser realizada a extração de características da mesma. Então, primeiramente, transformaram-se as imagens em escala cinza, baseando-se nos passos para processamento de imagens de Trier [23].

Para a extração de características das imagens da Base de Dados, escolheu-se os métodos de Limiarização e Projeção de Histogramas, com o auxílio da Biblioteca *OpenCV*. Esta biblioteca foi escolhida, pois “implementa uma variedade de ferramentas de interpretação de imagens, desde operações simples como um filtro de ruído, até operações complexas, tais como a análise de movimentos, reconhecimento de padrões e reconstrução em 3D” [15].

Com a limiarização foi possível definir o nível de limiar (*threshold*) a fim de se separar a parte de interesse da imagem com o fundo (Figura 3). As imagens foram processadas com um determinado nível de limiar, baseadas nas intensidades dos pixels dos histogramas, e posteriormente, o resultado foi observado. Caso algumas regiões de interesse estivessem sendo desconsideradas, este nível foi alterado. Este processo foi realizado até se obter o resultado esperado.



Figura 3. Limiarização aplicada à imagem da Base de Dados.

Utilizando a técnica de características estruturais, que consiste em “calcular alguns histogramas para formar o vetor de características”, segundo Cruz [24], a imagem foi normalizada para uma matriz de tamanho 200x200 pixels. Em seguida são calculadas as projeções de histogramas horizontal e vertical.

Para as definições a seguir, o valor de cada pixel é identificado mediante a  $m$ -ésima coluna e  $n$ -ésima linha da imagem, representado pela Equação (5) [24].

$$f(m, n) = a_{mn} \quad (5)$$

Em seguida, são calculadas as projeções horizontais e verticais de histograma. A projeção horizontal é calculada pela soma dos pixels pretos em cada linha da imagem, totalizando em 200 características. Esta projeção de histograma é dada pela Equação (6):

$$H_h(m) = \sum_n f(m, n) \quad (6)$$

A projeção vertical é calculada pela soma dos pixels pretos de cada coluna da imagem, totalizando em um vetor com 200 características. Esta projeção de histograma é definida pela Equação (7).

$$H_v(m) = \sum_n f(m, n) \quad (7)$$

Com as projeções de histogramas formadas, foi possível gerar um vetor de 200 posições para cada imagem, correspondente a cada configuração de mão. Este vetor, posteriormente, é utilizado para as entradas da Rede Kohonen.

## 6.2 Metodologia

Para o desenvolvimento do sistema foi escolhida a linguagem Python, por ser uma linguagem livre, moderna e com bibliotecas úteis ao desenvolvimento do trabalho. Essa linguagem “possui sintaxe elegante e tipagem dinâmica, em adição à sua natureza interpretada, o que a torna ideal para o desenvolvimento rápido de aplicações em diversas áreas e na maioria das plataformas, pois permite a escrita de programas compactos e legíveis” [24], e desta forma, a prototipação de soluções é rápida.

## 6.3 Treinamento da Kohonen

Para realizar o treinamento da rede Kohonen, é necessário especificar alguns parâmetros que definem a estrutura da rede e que controlam o treinamento propriamente dito. O principal objetivo é garantir um bom treinamento das configurações de mão, baseando-se na metodologia proposta do algoritmo da rede.

Foram utilizados para realizar o treinamento da rede: inicialização aleatória dos pesos, taxa de erro, apresentação dos vetores de entrada, número de épocas de treinamento, tipo de funções de vizinhança e aprendizado gaussianas, e formato retangular para o arranjo bidimensional (10x20) de neurônios em função dos vetores de entrada serem compostos de 200 nós.

Para o treinamento, foram separadas 70% das imagens, baseando-se nas recomendações existentes na literatura. Desta forma, criaram-se seis arquivos de treinos, nomeados “Treino1.txt, Treino2.txt, Treino3.txt” e assim por diante.

A Figura 4 mostra como é formada a Kohonen para a aplicação. Exemplifica-se com um vetor com 14 nós de entrada, e um mapa bidimensional de neurônios (2x5). Os nós de entrada correspondem às posições de um vetor de entrada, que é uma configuração de mão da LIBRAS (no caso de 200 posições). Cada arquivo de treino contém 61 linhas de vetores de características, formados pelas projeções de histogramas, correspondentes às 61 configurações escolhidas para serem trabalhadas. Cada um destes nós é ligado aos neurônios, e esta ligação contém um peso inicial aleatório, que será ajustado no decorrer do treinamento. Os neurônios estão dispostos de forma bidimensional (10x20). As imagens das configurações são identificadas pelo seu número.

Para esta identificação separou-se a primeira coluna de cada linha do arquivo.

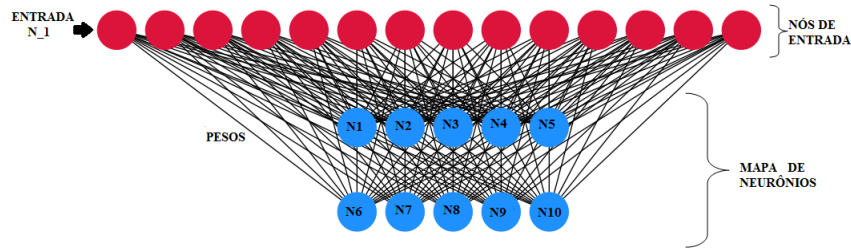


Figura 4. Imagem ilustrativa da Kohonen para a aplicação computacional.

Os resultados obtidos nos treinamentos foram salvos em arquivos separados. A implementação gera um arquivo compondo a matriz de pesos ajustada, que será utilizada para os testes da rede Kohonen. Também é gerado um arquivo demonstrando o processo de treinamento, ou seja, compõe-se de todos os neurônios que foram ativados para todas as classificações de mão.

O treinamento das configurações é realizado até o número de épocas ou a taxa de erro ser atingida, exibido na figura 5 o critério de parada. Esta taxa de erro é atingida se alguma “Distância Euclidiana for menor que a taxa de erro definida no sistema” [25]. O referido autor também indica que a taxa de erro pode ser no valor de 0.99.

```
#Inicia treinamento com
for iteracao in range (epocas):
    :
    :
    #Encontra menor distancia euclidiana
    menorDistancia();
    #Verifica o critério de parada
    if menorDistancia < taxa_de_erro:
        break
```

Figura 5. Trecho de código do critério de parada de treinamento.

Os resultados foram analisados mediante as taxas de acertos de cada classificação e índice Kappa. O Coeficiente Kappa pode ser definido como uma medida de associação usada para descrever e testar o grau de concordância (confiabilidade e precisão) na classificação[26]. Segundo Moreira [27], o valor resultante pelo cálculo do índice Kappa pode ser analisado mediante a qualidade da classificação explicitada na tabela 1.

*Tabela 1. Valor de referência do índice Kappa.*

Valor de kappa	Qualidade da classificação
< 0,00	Péssima
0,00 - 0,20	Ruim
0,20 - 0,40	Razoável
0,40 - 0,60	Boa
0,60 - 0,80	Muito boa
0,80 - 1,00	Excelente

## 7 Resultados e discussão

Foram separadas três pastas da Base de Dados para testar a metodologia de classificação proposta. Estas pastas são diferentes das utilizadas para o treinamento da rede Kohonen. Em cada pasta, realizou-se as etapas pré-processamento, segmentação das imagens e geração de arquivos com projeções de histogramas, compondo os vetores de entrada para testes. Estas etapas foram realizadas da mesma forma que os arquivos de treino.

Para obter facilidade na análise das classificações, optou-se que o sistema demonstrasse uma sequência de classificações, ou seja, cada configuração é listada com seu respectivo reconhecimento. Desta maneira, foi possível analisar os resultados apenas verificando quais classificações estavam corretas.

É possível analisar na tabela 2, a variação das taxas de aprendizado e vizinhança inicial, definidas conforme as equações 4 e 2, respectivamente, bem como as épocas para treinamento, que correspondem ao número de iterações para o processo de treinamento ser finalizado. Os acertos são referentes ao número de configurações de mão classificadas corretamente.

*Tabela 2. Resultados obtidos com as configurações de teste.*

Teste	Aprendizado	Vizinhança	Épocas	Acertos (%)
1º Teste	0,8	0.8	5	84
2º Teste	0,9	5	10	86
3º Teste	0,7	5	15	85
4º Teste	0,9	5	5	88
5º Teste	0,9	5	20	90
6º Teste	0,9	7	30	81



A taxa de vizinhança não foi variada muitas vezes. Segundo Haykin [6], o melhor valor para se iniciar a vizinhança da Kohonen, é estabelecer metade do diâmetro do mapa. O referido autor também indica que os valores de taxa de aprendizado se iniciem o mais próximo de “1”.

De acordo com os resultados apresentados na tabela 3, o valor calculado do índice Kappa é de 0,86. Conforme Moreira [27], apresentado na tabela 1, os resultados evidenciaram que a concordância pode ser considerada excelente. Este resultado apresentou que algumas configurações de mão semelhantes, torna mais difícil para o classificador discriminar as variações do atributo mensurado.

*Tabela 3. Resultado das classificações.*

Rótulo	T1	T2	T3	T4	T5	T6
0	0	0	0	15	0	0
1	14	1	1	1	1	36
2	2	2	2	37	2	2
3	3	39	3	3	3	7
4	4	4	4	4	4	4
5	5	5	5	5	5	5
6	6	6	6	6	6	6
7	7	7	7	7	7	7
8	8	8	8	8	8	8
9	9	9	9	9	9	9
10	12	11	47	10	10	10
11	11	11	11	11	11	11
12	32	12	12	12	39	12
13	13	13	6	13	13	43
14	14	14	14	60	6	14
15	15	15	15	15	15	15
16	16	16	16	16	16	16
17	17	24	51	17	17	17

<b>Rótulo</b>	<b>T1</b>	<b>T2</b>	<b>T3</b>	<b>T4</b>	<b>T5</b>	<b>T6</b>
18	18	18	40	28	18	56
19	41	19	19	19	19	19
20	20	23	24	20	55	20
21	21	21	21	21	38	21
22	60	22	22	22	22	22
23	23	23	23	23	23	23
24	24	24	24	24	24	24
25	25	25	25	25	25	25
26	49	26	26	26	26	39
27	57	27	27	27	27	27
28	28	43	28	28	28	58
29	29	29	29	29	29	29
30	34	30	30	37	30	39
31	31	31	31	31	31	31
32	32	32	32	32	51	32
33	33	8	55	33	33	33
34	34	34	34	34	34	34
35	35	35	35	35	35	35
36	36	36	36	36	36	36
37	37	50	37	37	37	37
38	38	38	38	38	38	38
39	39	39	45	57	39	39
40	60	40	40	40	40	40
41	41	41	8	41	41	41
42	42	42	42	42	42	54
43	43	43	43	43	43	43
44	44	58	9	44	44	44
45	45	45	45	45	45	45
46	53	46	46	46	46	47

Rótulo	T1	T2	T3	T4	T5	T6
47	47	47	47	47	47	47
48	48	48	48	48	48	48
49	49	49	49	49	49	49
50	50	50	50	50	50	7
51	51	51	51	51	51	51
52	52	52	52	52	52	52
53	53	53	53	53	9	53
54	54	54	54	54	54	54
55	55	55	55	55	55	55
56	56	56	56	56	56	56
57	57	57	57	57	57	57
58	58	58	58	58	58	58
59	59	59	59	59	59	59
60	60	60	60	60	60	60

## 8 Conclusões

Este trabalho apresentou classificação das configurações de mão da LIBRAS mediante treinamento de Rede Neural Artificial Kohonen, com as etapas de extração de características de imagens, aprendizado e testes.

A escolha da técnica para fazer a extração de características da imagem correspondeu ao esperado, pois foi possível obter a informação de interesse (configuração de mão) das imagens da Base de Dados, mediante o processo de limiarização. Porém, pela Base de Dados não ter sido produzida pelos autores, houve limitação com relação as etapas de processamento de imagens, tais como variação de métodos de aquisição e pré-processamento.

Considerando os resultados obtidos, nos testes de confiabilidade para classificação das configurações de mão, pode-se destacar os aspectos: o grau de concordância nos escores obtidos foi de 0,86 (concordância excelente); as configurações de mão com características semelhantes revelaram maior grau de confusão na classificação; houve

configurações que apresentaram 100% de acerto em todos os testes.

Estes dados demonstram evidências que a Kohonen apresenta confiabilidade para ser utilizado em classificação de configurações de mão, com técnicas de aprendizado eficientes para o estudo de caso.

Destaca-se que o presente trabalho não teve a pretensão de encerrar o assunto sobre aplicação de redes neurais artificiais em configurações de mão da LIBRAS e sim, divulgar a eficácia desta ferramenta para este tipo de problema.

## Referências

- [1] PISTORI, H. P.; MARTINS, P. S.; PEREIRA, M.C.; CASTRO JR, A. A.; JOSÉ NETO, J. Sigus – plataforma de apoio ao desenvolvimento de sistemas para inclusão digital de pessoas com necessidades especiais. In *IV Congresso Iberdiscap: Tecnologias de Apoio a Portadores de Deficiência*, Vitória, 2006.
- [2] BAPTISTA, F.F-Libras: Ambiente integrado de ensino – aprendizagem para a Língua Brasileira de Sinais. Dissertação de Mestrado. Centro de Ensino Eurípedes Soares da Rocha, Marília, 2007.
- [3] AJA. O que é LIBRAS. Disponível em: <http://www.libras.org.br/libras.php>  
Acessado em: mai/2012
- [4] GOLÇALVES, M.; NETTO, M.; COSTA, J. Explorando as Propriedades do Mapa Auto-organizável de Kohonen na Classificação de Imagens de Satélite. 2007. Disponível em: [http://www.cos.ufrj.br/~ines/enia07\\_html/pdf/27845.pdf](http://www.cos.ufrj.br/~ines/enia07_html/pdf/27845.pdf) Acessado em: set/2012
- [5] LUDWIG JR, O.; COSTA, E. M. M. Redes Neurais – Fundamentos e Aplicações com Programas em C. Rio de Janeiro: Editora Ciência Moderna Ltda, 2007.
- [6] HAYKIN, S. Redes Neurais: princípios e prática. 2.ed. Porto Alegre: Bookman, 2001.

- [7] TAFNER, M. S.; XERES, M.; RODRIGUES FILHO, I. W. Redes Neurais Artificiais: Introdução e Princípios de Neurocomputação. Blumenau: Editora da FURB, 1995.
- [8] REZENDE, S. O. (Org.). Sistemas Inteligentes: fundamentos e aplicações. São Paulo: Monole, 2005.
- [9] BRAGA, A. P.; CARVALHO, A. C. P. L. F.; LUDEMIR, T. B. Redes neurais artificiais: teoria e aplicações. Rio de Janeiro: LTC editora, 2000.
- [10] CARNEIRO, A. T. S.; CORTEZ, P. C.; COSTA, R. C. S. Reconhecimento dos gestos da LIBRAS com classificadores Neurais a partir dos Momentos invariantes de Hu, In *Interaction South-America*, São Paulo., p.190-195, 2009.
- [11] LAMAR, M. V. Reconhecimento de Posturas Manuais usando Redes Neurais. In *VI Congresso Brasileiro de Redes Neurais*, 2003, São Paulo : Centro Universitário Da Fei, p.207-212, 2003.
- [12] LOESCH, C. Redes Neurais artificiais: fundamentos e modelos. Blumenau: Editora da FURB, 1996.
- [13] FRANCISCO, C. Rede Kohonen: Uma ferramenta no estudo das relações tróficas entre espécies de peixes. 2004. Disponível em: <http://dspace.c3sl.ufpr.br/dspace/bitstream/handle/1884/637/Francisco,%20Cl%C3%A1udia%20Aparecida%20Cavalheiro.pdf?sequence=1> Acessado em: jun/2012.
- [14] CASTRO, F.; CASTRO, M. Mapas auto-organizáveis da Kohonen - SOM. 2001. Disponível em: [http://www.feng.pucrs.br/decastro/RNA\\_hp/pdf/RNA\\_C7.pdf](http://www.feng.pucrs.br/decastro/RNA_hp/pdf/RNA_C7.pdf) Acessado em: jun/2012.
- [15] MARENGONI, M.; STRINGHINI, D. Tutorial: Introdução à Visão Computacional usando OpenCV. Universidade Presbiteriana Mackenzie, 2009.
- [16] TRINDADE, F. H. V. Técnicas de Visão Computacional para Rastreamento de Olhar em Vídeos. 2009. Disponível em:

[http://almerindo.devin.com.br/index.php?option=com\\_content&view=article&id=78:tecnicas-de-computacao-visual-para-rastreamento-de-olhar-em-videos&catid=43:trabalhos-de-alunos&Itemid=18](http://almerindo.devin.com.br/index.php?option=com_content&view=article&id=78:tecnicas-de-computacao-visual-para-rastreamento-de-olhar-em-videos&catid=43:trabalhos-de-alunos&Itemid=18) Acessado em: maio/2012.

- [17] BASTOS, V. Segmentação. 2001. Disponível em: <http://paginas.ucpel.tche.br/vbastos/segmenta.htm> Acessado em: ago/2012.
- [18] MINDMANAGER, M. Limiarização. 2007. Disponível em: [http://www.gta.ufrj.br/grad/07\\_2/eliseu/Limiarizacao.html](http://www.gta.ufrj.br/grad/07_2/eliseu/Limiarizacao.html) Acessado em: ago/2012.
- [19] SIOLA, F. Desenvolvimento de um software para reconhecimento de sinais LIBRAS através de vídeo. Trabalho de Conclusão de Curso, Universidade Federal do ABC, Santo André, 2010.
- [20] STROBEL, K. L.; FERNANDES, S. Aspectos Linguísticos da LIBRAS. Secretaria do Estado da Educação, Superintendência de Educação, Departamento de Educação Especial. Curitiba, 1998.
- [21] OLIVEIRA, N. Os parâmetros da Libras. 2010. Disponível <http://www.guiame.com.br/noticias/ponto-de-vista/nilton-camara/os-parametros-da-libras.html> Acessado em: maio/2012
- [22] ANTUNES, D. Um modelo de descrição computacional da fonologia da Língua de Sinais Brasileira. Dissertação de Mestrado. Universidade Federal do Paraná, Curitiba, 2011.
- [23] TRIER, O. Feature extraction methods for character recognition – A survey. *Pattern Recogn*, v.19, n.4, p.641-662, 1996.
- [24] CRUZ, R. M. O. Análise de técnicas de extração de características para o reconhecimento de dígitos manuscritos. Trabalho de Conclusão de Curso. Universidade Federal de Pernambuco, Recife, 2008.
- [25] MAGRI, D. Redes Neurais Artificiais. 2011. Disponível em:

<http://www.slideshare.net/iaudesc/rna-redes-neurais-artificiais> Acessado em: ago/2012.

- [26] KOTZ, S.; JOHNSON, N. L. Encyclopedia of Statistical Sciences. New York: John Wiley & Sons; v.4, p.352-4. 1983.
- [27] MOREIRA, M. A. Fundamentos do Sensoriamento e Metodologia de Aplicação. 3.ed. Viçosa: Ed. UFV, 2011.